In many situations, we require to summarise the data by means of a simple number called a *descriptive measure*. A descriptive measure computed from the data of a *sample* is called a ***statistic***, while a descriptive measure computed from the data of a population is called a ***parameter***.

Of the various types of descriptive measures, we shall limit ourselves to only *two* important ones, namely

1. ***Measure of central tendency***
2. ***Measure of dispersion***

Measure of central tendency — There are *three* most commonly used measures of the central tendency of data. These are (i) the *mean*, (ii) the *median* and (iii) the *mode*. But what indeed are they?

15.2.1 Mean of a sample

Mean is simply the ***arithmetic mean*** of the data and is obtained by adding all the values in a sample or population and dividing the sum by the total number of values. It is also called ***sample mean*** or ***average***, in brief.

...ing corresponding frequencies

$$G = \sqrt[n]{(x_1)^{f_1}(x_2)^{f_2}\cdots(x_n)^{f_n}}$$

Harmonic mean (H.M.) of a set of values of a variate is the reciprocal of the arithmetic mean of the reciprocals of corresponding values of the variate.

$$H = \frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\cdots+\frac{1}{x_n}} = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}}$$

15.2.2 Median of a sample

The *median* is the measure of the central item when the values of variate are arranged in ascending or descending order.

The median of a finite set of values of a variate is the value that divides the set into two equal parts such that the number of values equal to or greater than the median is equal to the number of values equal to or less than the median.

If the number of values is *odd*, then the median is the *middle value* when the values are all arranged in order of magnitude. When the number is *even*, there are *two middle values* and the *median is the mean of the two middle values*. For instance, if the values are 18, 20, 3, 5, 25, 8, 16, then by arranging in ascending order we get 3, 5, 8, 16, 18, 20, 25. So, here $n = 7$ and $\frac{1}{2}(n+1)$th term, i.e., the 4th term is the median and the median $= 16$. If however the values are 20, 24, 6, 9, 15, 8, 13, 16, then by arranging in ascending order

we get $6, 8, 9, 13, 15, 16, 20, 24$. So, here $n = 8$ and the $(n/2)$th and the $[(n/2) + 1]$th i.e., the 4th and the 5th terms are the middle terms, corresponding to 13 and 15. So, the median $= (13 + 15)/2 = 14$.

The ***features of the median*** are (i) its *uniqueness*, there being only one median (or at best two) for a given set of data, (ii) its *simplicity*, in the sense that the median is easy to calculate and (iii) it is *not drastically affected by the extreme* or *atypical values*, unlike the mean.

15.2.3 Mode of a set of data

The mode of a set of data is that value of the variable that appears most frequently. Thus, it is the value of the maximum frequency or the point of greatest density. If the values in the set are all different, then plainly there is *no* mode. However, a set of values may have *more than one* mode.

For instance, the *mode* of the following distribution

$$3, 5, 14, 7, 9, 11, 13, 11, 18, 11, 22, 11, 19, 28, 11$$

is 11 which occurs 5 times and no other item occurs 5 times or more.

The following distribution however has *no* mode : $3, 5, 14, 7, 9, 11, 13, 18, 19, 28$, since the values of the set are all different.

The following set of values has *more than one* mode : $3, 3, 5, 7, 7, 11, 13, 7, 11, 19, 28, 11$ The modes are 7 and 11.

The mode, in general, is not of much importance. However, it is of great importance to the large scale manufacturer of consumption goods.

- Of the three measures of central tendency of data, the *mean has the greatest significance* and *the mode the least.* Mode is misleading when the distributions are small in number or highly non-symmetrical. Median however is helpful with very large or small values of distribution.
- Note that the *median always lies between the mean and the mode.* In fact,

$$\textbf{\textit{mean}} < \textbf{\textit{median}} < \textbf{\textit{mode}}$$

To compute the *variance* σ^2 *of a finite population* N, the procedure is the same as above, except division by N, and not by $N-1$.

$$\therefore \quad \sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{n} \qquad (\textit{finite population variance}) \tag{15.3.5}$$

where μ is the population mean.

Standard deviation — The *variance suffers from a drawback* in that it represents squared units and is thus not an appropriate measure of dispersion when we like to express the concept in terms of the original units.

To get a *measure of dispersion in original units*, we take simply the *square root of variance*, and thereby get what is called the *standard deviation* (S.D.). ***The standard deviation of a sample is the square root of its variance.*** Thus, by definition, the *standard deviation s of a sample of size n is given by*

$$s = \sqrt{s^2} = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}} \tag{15.3.6}$$

Similarly, *the standard deviation* σ *of a finite population of size* N *is given by*

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{N}} \tag{15.3.7}$$

Coefficient of variation — The standard deviation is undoubtedly an useful measure of variation within a given set of data. But *difficulty arises* when we like to compare the dispersion of two sets of data; in such a case, comparison of the two

[E] Some theorems.

THEOREM 1. **The law of total probability.**

If A and B are two mutually exclusive events, then

$$P(A \cup B) = P(A) + P(B)$$

or in general if $A_1, A_2, A_3, \ldots, A_n$ *are n mutually exclusive events, then*

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n).$$

Out of m mutually exclusive and equally likely cases, let a correspond to the occurrence of event A and b correspond to the occurrence of event B, so that the event $A \cup B$ occurs in $(a + b)$ cases due to the nature of the events A and B. Then,

$$P(A) = \frac{a}{m}, \quad P(B) = \frac{b}{m}$$

and

$$P(A \cup B) = \frac{a+b}{m}$$

$$= \frac{a}{m} + \frac{b}{m}$$

$$= P(A) + P(B). \qquad \ldots (1)$$

To generalize the theorem by mathematical induction, since the theorem is true for $n = 2$, let us assume that it is true for any positive integral value n, *i.e.*,

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_1) + \ldots + P(A_n). \qquad \ldots (2)$$

Let $A_1, A_2, A_3, \ldots, A_n, A_{n+1}$ be $(n + 1)$ mutually exclusive events, then

$$P(A_1 \cup A_2 \cup \ldots \cup A_n \cup A_{n+1}) = P(A_1 \cup A_2 \cup \ldots \cup A_n) + P(A_{n+1}) \text{ by (1)}$$

$$= P(A_1) + P(A_2) + \ldots + P(A_n) + P(A_{n+1}) \text{ by (2)}$$

This follows that the theorem is true for $n + 1$ and hence the theorem is true for any integral value n.

THEOREM 2. *For any two events A and B, the probability that either A or B or both occur is given by*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

$A \cup B$ can be decomposed into two mutually exclusive events $A - B$ and B, *i.e.*,

$$A \cup B = (A - B) \cup B.$$

where $\Sigma x_i = 1$
already occurred.

15.7 Probability distribution

The relationship between the values of a random variable and the probabilities of their occurrence can be summarised by a device called the ***probability distribution.*** It may be expressed in various forms — a table, a graph or a formula — and is a powerful tool for summarising and describing a set of data and arriving at scientific conclusions regarding a population of data on the basis of a sample of data drawn from the same population.

So, we may define ***probability distribution of a discrete variable as : a table or a graph or a formula and such other device so as to specify all its possible values along with their respective probabilities.***

Example — We give here an example. In a Health Camp conducted in a local High School, it was found that the prevalence of diseases among the students is as tabulated in the first two columns (next page).

The probability distribution of the discrete variable X where $X =$ no. of diseases in the population under study is shown in the third column; the values of X are $x_1 = 0,\ x_2 = 1, \cdots, x_8 = 7,\ x_9 = 8$. The probabilities have been computed by dividing their respective frequencies by the total frequency 1390. The values of $P(X = x)$ are all positive, less than 1 and their sum is 1, being the *characteristic of all probability distributions of discrete variables.*

Having the distribution, we can easily make probability statements about the random variable X. For instance, the probability that a randomly selected student having three diseases is $P(X = 3)$ and, from the third column of the table, is

Then from theorem 1, we have

$$P(A \cup B) = P[(A - B) \cup B]$$
$$= P(A - B) + P(B)$$

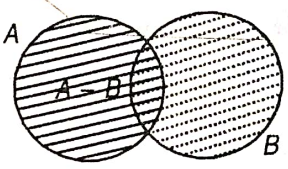

Fig. 15. 1

Now A can be decomposed into two mutually exclusive events $A - B$ and $A \cap B$, *i.e.*,

$$A = (A - B) \cup (A \cap B),$$

so that $P(A) = P(A - B) + P(A \cap B)$ by theorem 1

or $P(A - B) = P(A) - P(A \cap B)$.

Thus $P(A \cup B) = P(A) - P(A \cap B) + P(B)$

$$= P(A) + P(B) - P(A \cap B). \quad \ldots (1)$$

Aliter, $A \cup B = A \cup (B - A \cap B)$ where A and $B - A \cap B$ are disjoint.

$\therefore$ $P(A \cap B) = P(A) + P[B - A \cap B]$

we have $B = [A \cap B] \cup [B - A \cap B]$ where $A \cap B$ and $B - A \cap B$ are disjoint.

$\therefore$ $P(B) = P(A \cap B) + P[(B - A \cap B]$.

Hence $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

COROLLARY. For any events A, B, C applying this theorem twice we may easily get

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C)$$
$$- P(B \cap C) + P(A \cap B \cap C) \quad \ldots (2)$$

THEOREM 3. Let A be an event then its complementary event $\bar{A}$ will be disjoint to A, then $P(A) = 1 - P(\bar{A})$.

Since A, $\bar{A}$ are mutually exclusive events,

$$P(A \cup \bar{A}) = P(A) + P(\bar{A}).$$

But we know that $P(S) = 1 = P(A) + P(\bar{A})$

giving $P(A) = 1 - P(\bar{A})$.

THEOREM 4. If A and B are two events such that $A \subset B$, then

$$P(A) \leq P(B).$$

Since $A \subset B$, therefore B can be decomposed into two mutually exclusive events A and $B - A$, *i.e.*,

$$B = A \cup (B - A),$$

so that by Theorem 1,

$P(B) = P(A) + P(B - A)$

$\therefore P(B) \geq P(A)$ as $0 \leq P(B - A) \leq 1$.

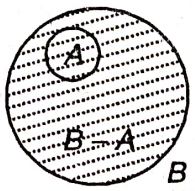

Fig. 15.2

THEOREM 5. If A and B are events, then

$$P(A \cup B) \leq P(A) + P(B).$$

By Theorem 2, we have

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$\therefore P(A \cup B) \leq P(A) + P(B)$ as $0 \leq P(A \cap B) \leq 1$.

COROLLARY. This result may be generalised in case of n events $A_1, A_2, A_3 \ldots, A_n$ as

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) \leq P(A_1) + P(A_2) + \ldots + P(A_n).$$

[F] Borel-field. We assume that a family B of certain subsets of S (outcome space) is given, which will be called events and that this family satisfies the following axioms :

B_1. The outcome space S and the empty set ϕ are in B, *i.e.*,

$$S \in B, \phi \in B.$$

• If however the n-events referred to are also exhaustive then $P = 1$ and the above result boils down to the note under the third elementary property in Art. 15.4.1.

• But what happens if the two events are *not mutually exclusive*?

If two events A and B are given, the probability that the event A, or the event B, or both occur is equal to the probability that A occurs plus the probability that B occurs minus the probability that the two events occur simultaneously.

Symbolically, $$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

15.6 Multiplication law of probability

If there are two independent events with known respective probabilities, then the probability that both will occur (i.e., joint occurrence) is the product of their respective individual probabilities.

Symbolically, $$P(A \cap B) = P(A)P(B)$$

Proof: Let A and B be two independent events and let A *occur* in m_1 ways and *fail to occur* in n_1 ways.

$$\therefore \quad P(A) = \frac{m_1}{m_1 + n_1}$$

Also, let B happen in m_2 ways and fail to occur in n_2 ways.

$$\therefore \quad P(B) = \frac{m_2}{m_2 + n_2}$$

Now, there are in all *four* possibilities which are as under.

1. A and B both happen : No. of ways $= m_1 m_2$
2. A happens, but B fails : No. of ways $= m_1 n_2$
3. A fails, but B happens : No. of ways $= n_1 m_2$
4. A and B both fail : No. of ways $= n_1 n_2$.

$\therefore$ Total number of ways $= m_1 m_2 + m_1 n_2 + n_1 m_2 + n_1 n_2$

$$= (m_1 + n_1)(m_2 + n_2)$$

$\therefore$ Probability that both the events A and B will occur is given by

$$P(A \cap B) = \frac{m_1 m_2}{(m_1 + n_1)(m_2 + n_2)}$$

$$= \left(\frac{m_1}{m_1 + n_1}\right) \times \left(\frac{m_2}{m_2 + n_2}\right)$$

$$= P(A) \times P(B)$$

• **Independent events** — If, of the two events A and B, the probability of event A is the same regardless of whether or not B occurs, then A and B are said to be *independent events.*

Therefore, chance of throwing a 7 with two dice $= \frac{6}{36} = \frac{1}{6}$.

If $e_1, e_2, e_3, e_4, e_5, e_6$ denote the events of throwing a seven as (1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1) respectively, then the event $A = \{e_1, e_2, e_3, e_4, e_5, e_6\}$, so that $P(A) = P(e_1) + P(e_2) + P(e_3) + P(e_4) + P(e_5) + P(e_6)$

But $$P(e_1) = P(e_2) = \ldots = P(e_6) = \frac{1}{36}.$$

$$\therefore \quad P(A) = \frac{1}{36} + \frac{1}{26} + \ldots \text{times}.$$

$$= \frac{1}{6}.$$

Problem 31. *For any two events A and B prove that*

$$P[(A \cap \bar{B}) \cup (B \cap \bar{A})] = P(A) + P(B) - 2P(A \cap B)$$

A can be decomposed into two mutually exclusive events $A \cap \bar{B}$ and $A \cap B$, *i.e.*,

$$A = (A \cap \bar{B}) \cup (A \cap B)$$

Similarly $$B = (\bar{A} \cap B) \cup (A \cap B)$$

$$\therefore \quad P(A) = P(A \cap \bar{B}) + P(A \cap B)$$

and $$P(B) = P(\bar{A} \cap B) + P(A \cap B).$$

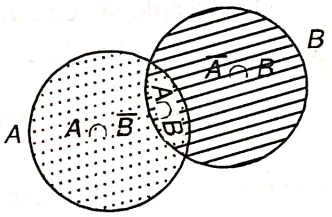

Fig. 15.3

Adding the last two results,

$$P(A) + P(B) = P(A \cap \bar{B}) + P(\bar{A} \cap B) + 2P(A \cap B)$$

$$= P[(A \cap \bar{B}) \cup (B \cap \bar{A})] + 2P(A \cap B)$$

as $A \cap \bar{B}$ and $B \cap \bar{A}$ are mutually exclusive events,

$\therefore P[(A \cap \bar{B}) \cup (B \cap \bar{A})] = P(A) + P(B) - 2P(A \cap B)$.

Problem 32. *From a pack of cards one card is selected at random. What is the probability that the card is a spade, an honour or a deuce ?*

Denoting by A, B, C respectively the events of drawing a spade, an honour and a deuce, we have

$$P(A) = \frac{{}^{13}C_1}{{}^{52}C_1} = \frac{13}{52}, \quad P(B) = \frac{{}^{12}C_1}{{}^{52}C_1} = \frac{12}{52}, \quad P(C) = \frac{{}^{4}C_1}{{}^{52}C_1} = \frac{4}{52}$$

$$P(A \cap B) = \frac{{}^{3}C_1}{{}^{52}C_1} = \frac{3}{52}, \quad P(A \cap C) = \frac{{}^{1}C_1}{{}^{52}C_1} = \frac{1}{52}, \quad P(B \cap C) = 0$$

and $$P(A \cap B \cap C) = 0,$$

Required probability

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

$$= \frac{13}{52} + \frac{12}{52} + \frac{4}{52} - \frac{3}{52} - \frac{1}{52} - 0 + 0 = \frac{25}{32}.$$

Problem 33. *An integer is chosen at random from the first 200 positive integers. What is the probability that the integer chosen is divisible by 6 or by 8?*

Let A denote the event that an integer selected is dividible by 6 and B denote the event that an integer selected is divisible by 8. Then the sample space is

• Note carefully that the terms '*mutually exclusive*' or '*independent*' do not convey the same meaning.

• **Complementary events** — Two events A and $\bar{A}$ are said to be ***complementary*** *if* the probability of event A is equal to 1 *minus* the probability of event $\bar{A}$, called the ***complement of A***, i.e., *if*

$$P(A) = 1 - P(\bar{A}) \;\Rightarrow\; P(\bar{A}) = 1 - P(A)$$

This immediately follows from the *second elementary property* of probability, for the event A and its complement $\bar{A}$, are *mutually exclusive*.

• The multiplication law may be *generalized* for n-number of *independent* events $E_1, E_2, \cdots, E_n$ and the probability P of their simultaneous happening is

$$P(E_1\, E_2 \cdots E_n) = P(E_1)P(E_2) \cdots P(E_n)$$

• If two events are *dependent*, then we shall have

$$P(E_1 \cup E_2) = P(E_1) \times P(E_2/E_1)$$

where E_2/E_1 denotes the *conditional probability* of E_2, on the supposition that E_1 has already occurred.

15.7 Probability distribution

lues of a random variable and the probabilities of their
the ***probability distribution***. It may
nd is a powerful

$$\therefore \quad P(A) = P(A_1) + P(A_2) + P(A_3)$$
$$= \frac{{}^3C_1 \times {}^6C_2 + {}^3C_2 \times {}^6C_1 \times {}^3C_3 \times {}^6C_0}{{}^9C_3}$$
$$= \frac{16}{21}.$$

Similarly, $P(B)$ can be found to be $\frac{1}{3}$ and hence the required ratio.

[H] Conditional probability. *If A and B are two dependent events in the sample space S, then the conditional probability of A given B (i.e., the probability of occurring of the event A on the assumption that B has already occurred) is defined as*

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \qquad \ldots (1)$$

provided $P(B) \neq 0$.

Similarly the conditional probability of B given A is defined as

$$P(B/A) = \frac{P(B \cap A)}{P(A)} \qquad \ldots (2)$$

provided $P(A) \neq 0$.

Note. In the first case if $P(B) = 0$ or in the second case $P(A) = 0$, the conditional probability is undefined.

[I] The law of compound probability. In the previous article we have defined the conditional probability of A given B as

$$P(A/B) = \frac{P(A \cap B)}{P(B)}.$$

Multiplying both sides by $P(B)$, this becomes

$$P(B)\, P(A/B) = P(A \cap B),$$

i.e., $$P(A \cap B) = P(B)\, P(A/B).$$

Similarly $$P(B \cap A) = P(A)\, P(B/A).$$

But $$B \cap A = A \cap B.$$

$$\therefore \quad P(B \cap A) = P(A \cap B),$$

so that $$P(A \cap B) = P(A).\, P(B/A) = P(B) \cdot P(A/B).$$

This gives the law of compound probability.

COROLLARY 1. In case of three events A, B, C, this law becomes

$$P(A \cap B \cap C) = P(A).P(B/A).\, P(C/A \cap B).$$

COROLLARY 2. In case of n (≥ 2) events $A_1, A_2, A_3, \ldots A_n$ for which $P(A_1 \cap A_2 \cap \ldots \cap A_n) \neq 0$, then the compound probability

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1).P(A_2/A_1).P(A_3/A_1 \cap A_2) \ldots P[A_n/A_1 \cap A_2 \cap \ldots \cap A_{n-1}].$$

This may be proved by the method of induction as follows:

Denoting by S_n the statement given above and denoting by N the set of those integers n for which S_n is true, we have by the method of induction for $N > 1$,

(*i*) $2 \in N$ for S_2 is the statement which is true,

(*ii*) $k \in N$, k being assumed an integer > 1.

► **Example 15.19** *X is a continuous random variable with probability density function given by*

$$\begin{aligned} f(x) &= kx, && (0 \le x < 2) \\ &= 2k, && (2 \le x < 4) \\ &= -kx + 6k, && (4 \le x < 6) \end{aligned}$$

Find the value of k and also the mean value of X. (JNTU)

Solution. Since the total probability must be unity, we must have

$$\int_0^6 f(x)dx = 1$$

$$\Rightarrow \int_0^2 kx\,dx + \int_2^4 2k\,dx + \int_4^6 (-kx + 6k)dx = 1$$

$$\Rightarrow k\left|\frac{x^2}{2}\right|_0^2 + 2k\left|x\right|_2^4 + \left[-kx^2/2 + 6kx\right]_4^6 = 1$$

$$\Rightarrow 2k + 4k + (-10k + 12k) = 1 \quad \Rightarrow \quad k = 1/8$$

$$\begin{aligned} \text{Mean of } X &= \int_0^6 xf(x)dx \\ &= \int_0^2 kx^2dx + \int_2^4 2kx\,dx + \int_4^6 x(-kx + 6k)dx \\ &= k\left|x^3/3\right|_0^2 + 2k\left|x^2/2\right|_2^4 + \left[-k\left|x^3/3\right|_4^6 + 6k\left|x^2/2\right|_4^6\right] \\ &= k(8/3) + k(12) - k(152/3) + 6k(10) \\ &= \frac{1}{8} \times \left[\frac{8}{3} + 12 - \frac{152}{3} + 60\right] \\ &= \frac{1}{8} \times 24 = \mathbf{3}. \end{aligned}$$

► **Example 15.20** *Deduce the exponential law of radioactive decay purely from the probability consideration.*

Solution. We start with the only assumption that the probability p for an atom to disintegrate in a time interval Δt is independent of the past history of the atom and the same for all atoms of the same type.

Probability of decay, for sufficiently small intervals, is proportional to Δt.

$$\therefore \quad p = \lambda \Delta t,$$

the element

If we prove that $k + 1 \in N$, then the statement becomes universally true.

For S_k the statement is

$$P(A_1 \cap A_2 \cap \ldots \cap A_k) = P(A_1)\, P(A_2/A) \ldots P(A_k/A_1 \cap A_2 \cap \ldots \cap A_{k-1}). \quad \ldots (1)$$

It may be verified by using the definition of conditional probability and properties of set-intersection, that

$$\frac{P(A_1 \cap A_2 \cap \ldots \cap A_{k+1})}{P(A_1 \cap A_2 \cap \ldots \cap A_k)} = P(A_{k+1} / A_1 \cap A_2 \cap \ldots \cap A_k). \quad \ldots (2)$$

Multiplying (1) and (2) together, we get

$$P(A_1 \cap A_2 \cap \ldots \cap A_{k+1}) = P(A_1).P(A_2/A_1) \ldots P(A_{k+1}/A_1 \cap A_2 \cap \ldots \cap A_k),$$

which shows that if $k \in N$, then $k + 1 \in N$, *i.e.*, the statement is true for all values of k and hence the result holds good for any integral value n.

COROLLARY 3. Since B can be decomposed into mutually exclusive events $A \cap B$ and $\bar{A} \cap B$, *i.e.*,

$$B = (A \cap B) \cup (\bar{A} \cap B)$$

we have,
$$\begin{aligned} P(B) &= P[(A \cap B) \cup (\bar{A} \cap B)] \\ &= P(A \cap B) + P(\bar{A} \cap B) \text{ as } P(E \cup F) = F(E) + P(F) \\ &= P(A).P(B/A) + P(\bar{A})\ P(B/A) \end{aligned}$$

Problem 36. *Prove the following cases assuming that in each case the conditional probabilities are defined.*

(i) For any event A, $0 \le P(A/B) \le 1$, where $P(B) > 0$, B being another event.

(ii) For a certain event S, $P(S/B) = 1$, where $P(B) > 0$.

(iii) $P(A_1 \cup A_2/B) = P(A_1/B) + P(A_2/B) - P(A_1 \cap A_2/B)$.

But A_1 and A_2 are mutually exclusive and $P(B) > 0$,

$$P(A_1 \cup A_2/B) = P(A_1/B) + P(A_2/B)$$

or in general if $A_1, A_2, \ldots A_n$ is a sequence of n mutually exclusive events, then

$$P(A_1 \cup A_2 \cup \ldots \cup A_n/B) = (A_1/B) + P(A_2/B) + \ldots + P(A_n/B)$$

(iv) If A and B are mutually exclusive and $P(A \cup B) \neq 0$; then

$$P(A/A \cup B) = \frac{P(A)}{P(A) + P(B)}.$$

(*i*) We have $A \cap B \subset B$, so that B can be decomposed into two mutually exclusive events $A \cap B$ and $B - A \cap B$, *i.e.*,

$$B = (A \cap B) \cup (B - A \cap B)$$

$$\therefore \quad P(B) = P(A \cap B) + P(B - A \cap B)$$

$$\geq P(A \cap B) \text{ as } P(B - A \cap B) \geq 0$$

or $\quad P(A \cap B) \le P(B).$

Thus, $\quad P(A/B) = \dfrac{P(A \cap B)}{P(B)} \le 1.$

Also $\quad P(A/B) \geq 0.$

$\therefore \quad 0 \le P(A/B) \le 1.$

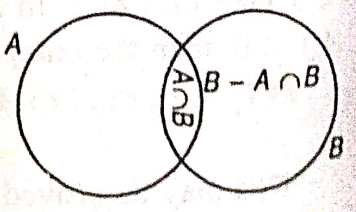

Fig. 15.4

(*iv*) Since A and B are mutually exclusive events, we have

$$A \cup B = \phi$$

and $$A \cap (A \cup B) = (A \cup A) \cap (A \cup B)$$
$$= A \cap \phi = A$$

so that $$P[A \cap (A \cup B)] = P(A)$$

and also $$P(A \cup B) = P(A) + P(B)$$

Hence $$P(A/A \cup B) = \frac{P[A \cap (A \cup B)]}{P(A \cup B)}$$
$$= \frac{P(A)}{P(A) + P(B)}$$

Problem 37. *If A and B are two events such that $P(A) = \frac{3}{8}, P(B) = \frac{5}{8}$ and $P(A \cup B) = \frac{3}{4}$, find $P(A/B)$ and $P(B/A)$.*

We have $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

i.e., $$\frac{3}{4} = \frac{3}{8} + \frac{5}{8} - P(A \cap B)$$

giving $$P(A \cap B) = \frac{1}{4}$$

Now $$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{4}}{\frac{5}{8}} = \frac{2}{5}$$

and $$P(B/A) = \frac{P(B \cap A)}{P(A)} = \frac{\frac{1}{4}}{\frac{5}{8}} = \frac{2}{5}$$ as $B \cap A = A \cap B$.

Problem 38. *If A and B are two events such that $P(A) = \frac{1}{3}, P(B) = \frac{1}{4}$ and $P(A \cup B) = \frac{1}{2}$, find $P(A/B)$, $P(B/A)$, $P(A \cap \bar{B})$ and $P(A/\bar{B})$.*

We have $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

i.e., $$\frac{1}{2} = \frac{1}{3} + \frac{1}{4} - P(A \cap B)$$ giving $P(A \cap B) = \frac{1}{12}$

$\therefore$ $$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{12}}{\frac{1}{4}} = \frac{1}{3}$$

$$P(B/A) = \frac{P(B \cap A)}{P(A)} = \frac{\frac{1}{12}}{\frac{1}{3}} = \frac{1}{4}.$$

Now, $$P(\bar{B}) = 1 - P(B) = 1 - \frac{1}{4} = \frac{3}{4}.$$

A can be decomposed into two mutually exclusive events $A \cap B$ and $A \cap \bar{B}$, so that

$$A = (A \cap B) \cup (A \cup \bar{B})$$

$$\therefore \quad P(A) = P[(A \cap B) \cup (A \cap \bar{B})]$$

$$= P(A \cap B) + P(A \cap \bar{B}),$$

so that $$P(A \cap \bar{B}) = P(A) - P(A \cap B)$$

$$= \frac{1}{3} - \frac{1}{12} = \frac{1}{4}$$

also $$P(A/\bar{B}) = \frac{P(A \cap \bar{B})}{P(\bar{B})} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}.$$

Problem 39. *A bag contains 3 block and 4 white balls. Two balls are drawn at random one at a time without replacement.*

(i) What is the probability that the second ball selected is white?

(ii) What is the conditional probability that the first ball selected is white if the second ball is known to be white?

Let W_1, W_2 dentoe the events of drawing a white ball in first and second draw, then

$$P(W_1) = \frac{{}^2C_1}{{}^7C_1} = \frac{4}{7}, \quad P(W_2) = \frac{{}^4C_1}{{}^7C_1} = \frac{4}{7}$$

$$\therefore \quad P(\overline{W_1}) = 1 - P(W_1) = 1 - \frac{4}{7} = \frac{3}{7}$$

$$P(W_2/W_1) = \frac{{}^3C_1}{{}^6C_1} = \frac{3}{6} = \frac{1}{2}$$

$$P(W_2/\overline{W_1}) = \frac{{}^4C_1}{{}^6C_1} = \frac{4}{6} = \frac{2}{3}.$$

Also $$P(W_2/W_1) = \frac{P(W_2 \cap W_1)}{P(W_1)}$$

i.e., $$\frac{1}{2} = \frac{P(W_2 \cap W_1)}{\frac{4}{7}}$$

or $$P(W_2 \cap W_1) = \frac{4}{7} \cdot \frac{1}{2} = \frac{2}{7}.$$

(i) $$P(W_2) = P[(W_1 \cap W_2) \cup (\bar{W}_1 \cap W_2)]$$

$$= P(W_1 \cap W_2) + P(\bar{W}_1 \cap W_2)$$

$$= P(W_1) \cdot P(W_2/W_1) + P(\bar{W}_1) \cdot P(W_2/\bar{W}_1)$$

$$= \frac{4}{7} \cdot \frac{1}{2} + \frac{3}{7} \cdot \frac{2}{3} = \frac{2}{7} + \frac{2}{7} = \frac{4}{7}.$$

(ii) $$P(W_1/W_2) = \frac{P(W_1 \cap W_2)}{P(W_2)}$$

$$= \frac{\frac{2}{7}}{\frac{4}{7}} = \frac{1}{2}.$$

Applying the result of the conditional probability, *i.e.*,

$$P(A \mid A_j) = \frac{P(A \cap A_j)}{P(A_j)} \text{ giving } P(A \cap A_j)$$

$$= P(A_j).\, P(A \mid A_j)$$

we have $$P(A) = \sum_{j=1}^{n} P(A_j).\, P(A \mid A_j), \qquad \ldots (1)$$

which proves the lemma.

Statement of Baye's theorem. *If an event A can occur only if one of the mutually exclusive events $A_1, A_2, \ldots, A_n$, i.e., $A \subset \bigcup_{k=1}^{k} A_k$, $A_k \cap A_j = \phi$ when $k \neq j$ and suppose we are given the probabilities $P(A_k)$, $k = 1, 2, \ldots, n$ and the conditional probabilities $P(A \mid A_k)$, then we are required to find the probability of A_k when it is given that A has already occurred and $P(A) > 0$ for each integer k $(1 \leq k \leq n)$ then Baye's formula is*

$$P(A_k \mid A) = \frac{P(A_k)\, P(A \mid A_k)}{\sum_{j=1}^{n} P(A_j)\, P(A \mid A_j)}.$$

Proof. By the definition of conditional probability, we have

$$P(A_k \mid A) = \frac{P(A \cap A_k)}{P(A)}$$

i.e., $$P(A \cap A_k) = P(A)\, P(A_k) \mid A)$$

$$= P(A_k)\, P(A \mid A_k)$$

and by (1). $$P(A) = \sum_{j=1}^{n} P(A_j).P(A \mid A_j).$$

$\therefore$ $$P(A_k/A) = \frac{P(A_k).P(A/A_k)}{\sum_{j=1}^{n} P(A_j).P(A/A_j)}.$$

Problem 40. *There are three coins, identical in appearance, one of which is ideal and the other two biased with probabilities $\frac{1}{3}$ and $\frac{2}{3}$ respectively for a head. One coin is taken at random and tossed twice. If a head appears both the times, what is the probability that the ideal coin was chosen.*

Let A_1, A_2, A_3 denote the events of choosing the 1st (ideal), 2nd and 3rd coins respectively. Then

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}.$$

Let A be the event of obtaining 2 heads in two tosses of the selected coin.

Probability of getting a head in a toss $= \frac{1}{2}$.

$\therefore$ $$P(A/A_1) = \left(\frac{1}{2}\right)^2 = \frac{1}{4}.$$

Probability of turning head up with 2nd coin is $\frac{1}{3}$ and that with 3rd coin is $\frac{2}{3}$; therefore in two tosses,

$$P(A/A_2) = \left(\frac{1}{3}\right)^2 = \frac{1}{9} \text{ and } P(A/A_3) = \left(\frac{2}{3}\right)^2 = \frac{4}{9}$$

Using Baye's formula,

$$P(A_1/A) = \frac{P(A_1)\,P(A/A_1)}{\sum_{j=1}^{3} P(A_j).P(A/A_j)}$$

$$= \frac{\frac{1}{3}\cdot\frac{1}{4}}{\frac{1}{3}\cdot\frac{1}{4}+\frac{1}{3}\cdot\frac{1}{9}+\frac{1}{3}\cdot\frac{4}{9}}$$

$$= \frac{\frac{1}{12}}{\frac{1}{12}+\frac{1}{27}+\frac{4}{27}} = \frac{\frac{1}{12}}{\frac{29}{108}} = \frac{1}{12}\times\frac{108}{29} = \frac{9}{29}.$$

Problem 41. *In a bolt factory, machines A, B, C manufacture respectively 25, 35 and 40% of the total. Of their output, 5, 4 and 2% are defective bolts. A bolt is drawn random from the procedure and is found defective. What are the probabilities that it was manufactured by machine A, B or C?*

Here $P(A) = \frac{25}{100}$, $P(B) = \frac{35}{100}$, $P(C) = \frac{40}{100}$.

Let E denote the event of drawing a defective bolt.

$$P(E/A) = \frac{5}{100}, P(E/B) = \frac{4}{100}, P(E/C) = \frac{2}{100}.$$

Using Baye's formula

$$P(A/E) = \frac{P(A)\;P(E\mid A)}{P(A)\,P(E\mid A)+P(B)\;P(E\mid B)+P(C)\,P(E\mid C)}$$

$$= \frac{\frac{25}{100}\times\frac{5}{100}}{\frac{25}{100}\times\frac{5}{100}+\frac{35}{100}\times\frac{4}{100}+\frac{40}{100}\times\frac{2}{100}}$$

$$= \frac{125}{125+140+80} = \frac{125}{345}.$$

Similarly $P(B \mid E) = \frac{140}{345}$ and $P(C \mid E) = \frac{80}{345}$.

[L] Independent events. *The two events A and B are said to be stochastically independent if and only if $P(A \mid B) = P(A)$, i.e., the happening of the event B does not affect the happening of A.*

We have

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \text{ and } P(B \mid A) = \frac{P(B \cap A)}{P(A)}$$

The two events are independent of each other when

$$P(A \mid B) = P(A),\ P(A) > 0,$$
$$P(B \mid A) = P(B),\ P(B) > 0.$$

$\therefore$ the chance of exactly r successes in n trials $= {}^nC_r p^r q^{n-r}$.

Putting $r = 1, 2, 3,...$ in succession we get the probabilities of exctly 1, 2, 3,...successes.

But we know that ${}^nC_r p^r q^{n-r}$ is $(r+1)$th term in the binomial expansion of $(q+p)^n$; *hence the probability that the event will happen exactly r times in n trials is the* $(r+1)$*th term in the binomial expansion of* $(q+p)^n$.

[B] Multinomial theorem. *If there are n dice with faces marked from 1 to f and these are thrown at random, then we have to find the chance that the sum of the numbers exhibited on the uppermost faces is equal to p.*

Since a die has f faces, it can fall in f ways. Similarly the second die can fall in f ways. Thus the two dice can fall together in $f \times f$, *i.e.*, f^2 ways. Similarly we can show that when the n dice fall together, they can fall in f^n ways.

Now the number of ways in which the numbers thrown will have p for their sum is equal to the coefficient of x^p in the expansion of

$$(x^1 + x^2 + x^3 + ... + x^f)^n$$

because this coefficient arises out of the different ways in which n of the indices 1, 2, 3,..., f can be taken so as to form p by addition.

Hence the required probability

$$= \frac{\text{Coeff. of } x^p \text{ in } (x^1 + x^2 + ... + x^f)^n}{f^n}$$

$$= \frac{\text{Coeff. of } x^p \text{ in } x^n \left[\frac{1-x^f}{1-x}\right]^n}{f^n}$$

$$= \frac{\text{Coeff. of } x^{p-n} \text{ in } (1-x^f)^n\,(1-x)^{-n}}{f^n}$$

Problem 44. *What is the chance that a person with two dice (the faces of each being numbered 1 to 6) will throw aces exactly 4 times in 6 trials.*

Probability of throwing an ace with a die $= \frac{1}{6}$.

$\therefore$ Probability of throwing an ace with two dice in a single throw

$$= \frac{1}{6}\times\frac{1}{6} = \frac{1}{36}.$$

so that the probability of not throwing an ace in a throw

$$= 1 - \frac{1}{36} = \frac{35}{36}.$$

The required chance of throwing exactly four aces in 6 trials is (4 + 1)th, *i.e.*, 5th term in the expansion of $\left(\frac{35}{36}+\frac{1}{36}\right)^6$.

$\therefore$ the required chance $= {}^6C_4\left(\frac{35}{36}\right)^2\left(\frac{1}{36}\right)^4$.

Problem 45. *An experiment succeeds twice as often as it fails. Find the chance that in the next six trials there will be at least 4 successes.*

Let p the probability of success and q that of failure. Then

$$p = 2q \quad \text{and} \quad p + q = 1$$

which give $p = \frac{2}{3}, q = \frac{1}{3}$.

The probabilities of 4, 5 and 6 successes are the 5th, 6th and 7th terms in the expansion of $\left(\frac{1}{3}+\frac{2}{3}\right)^6$, *i.e.*, they are ${}^6C_4\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^4$, ${}^6C_5\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^5$ and ${}^6C_6\left(\frac{2}{3}\right)^6$.

The three events being mutually exclusive, the required chance

$$= {}^6C_4\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^4 + {}^6C_5\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^5 + {}^6C_6\left(\frac{1}{3}\right)^6$$

$$= \frac{2^4}{729}(15 + 6\times 2 + 4) = \frac{496}{729}.$$

Problem 46. *An ordinary six-faced die is thrown 4 times. What are the probabilities of obtaining 4, 3, 2, 1, 0 aces ?*

Lst p be the chance of obtaining an ace in a single throw and q that of not obtaining an ace. Then

$$p = \frac{1}{6} \text{ and } q = 1 - \frac{1}{6} = \frac{5}{6}.$$

When a die is thrown 4 times, the chances of 4, 3, 2, 1, 0 aces are the 5th, 4th, 3rd, 2nd and 1st terms in the binomial expansion of $(q + p)^4$, *i.e.*, $\left(\frac{5}{6}+\frac{1}{6}\right)^4$.

$\therefore$ Probability of 4 aces $= {}^4C_4\left(\frac{1}{6}\right)^4 = \frac{1}{1296}$.

$\therefore$ Probability of 3 aces $= {}^4C_3\left(\frac{5}{6}\right)\left(\frac{1}{6}\right)^3 = \frac{20}{1296}$.

$\therefore$ Probability of 2 aces $= {}^4C_2\left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right)^2 = \frac{150}{1296}$.

Note. The curve given by $y = y_0 e^{-x^2/2\sigma^2}$ is said to be the normal curve.

A **normal curve** *is symmetrical about the point $x = 0$ where the ordinate has its maximum value. In a normal curve, the mean, the median and mode coincide.*

[C] The Normal Distribution

We have just introduced that the equation to the normal curve is

$$y = y_0 e^{-x^2/2\sigma^2} \quad \ldots (1)$$

Its area $= \int_{-\infty}^{\infty} y_0 e^{-x^2/2\sigma^2} dx$

$= 2y_0 \int_0^{\infty} e^{-x^2/2\sigma^2} dx$, being symmetrical

$= 2y_0 \cdot \dfrac{\sqrt{\pi}}{2 \cdot \dfrac{1}{\sqrt{2\sigma}}}$ as $\int_0^{\infty} e^{-a^2x^2} dx = \dfrac{\sqrt{\pi}}{2a}$ if $a > 0$

$= \sigma y_0 \sqrt{(2\pi)}$ ... (2)

$= 2.506627\, y_0\sigma$ by putting the value of $\sqrt{(2\pi)}$.

In order to make (1), the normal probability curve, the value of y_0 be so determined that the total frequency is one, *i.e.*, the area of the normal curve is unity and this will be so if from (2)

$$\sigma y_0 \sqrt{(2\pi)} = 1$$

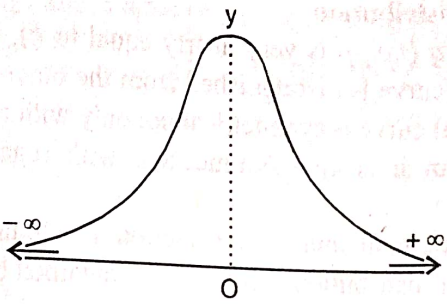

Fig. 15.5 (b)

i.e.,
$$y_0 = \frac{1}{\sigma\sqrt{(2\pi)}} \quad \ldots (3)$$

Substituting this value of y_0 in (1), the standard form of the normal curve becomes

$$y = \frac{1}{\sigma\sqrt{(2\pi)}} e^{-x^2/2\sigma^2} \quad \ldots (4)$$

which is the normal distribution.

In deriving the form (1) of normal curve, we have taken the mean at the origin, but if however we take another point as the origin such that the excess of the mean over the arbitrary origins is m, the form of the normal curve is

$$y = \frac{1}{\sigma\sqrt{(2\pi)}} e^{-(x-m)^2/2\sigma^2} \quad \ldots (5)$$

which represents the standard form of the normal curve with origin at $(m, 0)$.

Physical conditions leading a normal curve

(*i*) The casual forces that affect individual events are mutually independent.

(*ii*) The casual forces of equal magnitude are very large in number.

(*iii*) The casual forces operate in such a way that the maximum frequencies are clustered around the mean value thereby giving a symmetrical curve. Conclusively the deviations below the mean are equal in number and magnitude to those above the mean.

(*iv*) The normal distribution can be used as an error distribution by inquiring what law of distribution errors of observation should obey in order to make the mean of a set of measurements the most probable value of the 'true' magnitude.

Hence according to Gauss, if we call the '**precision**' h, such that $h^2 = \frac{1}{2\sigma^2}$, the form (4) becomes

$$y = \frac{h}{\sqrt{\pi}} e^{-h^2x^2} \quad \ldots (6)$$

It is clear form (6) that as h increases, the normal curve would become narrower and as such h is a measure of closeness of the mass of observations to the true value.

Definition of a normal distribution. *A normal distribution is a continuous distribution given by*

$$y = \frac{1}{\sigma\sqrt{(2\pi)}} e^{-\frac{1}{2}[(x-m)/\sigma]^2}$$

where x is a continuous normal variate distributed with probability density function $f(x) = \frac{1}{\sigma\sqrt{(2\pi)}} e^{-\frac{1}{2}[(x-m)/\sigma]^2}$, *with mean m and standard deviation* σ.